J. WESLEY HUTCHINSON, JOSEPH W. ALBA, and ERIC M. EISENSTEIN*

Managers use numerical data as the basis for many decisions. This research investigates how data on prior advertising expenditures and sales outcomes are used in budget allocation decisions and attempts to answer three important questions about data-based inferences. First, do biases exist that are strong enough to lead to seriously suboptimal decisions? Second, do graphical data displays, real-world experience, or explicit training reduce any observed biases? Third, are the observed biases well explained by a relatively small set of natural heuristics that managers use when making data-based allocation decisions? The results suggest answers of yes, no, and yes, respectively. The authors identify three broad classes of heuristics: difference-based (which assess causation by comparing adjacent changes in expenditures to changes in sales), trend-based (which assess causation by comparing overall trends in expenditures and sales), and exemplar-based (which emulate the allocation pattern of the observations with the highest sales). All three heuristics create biases in some situations. Overall, exemplar-based heuristics were used most frequently and induced the greatest biasing of the three (sometimes allocating the most to an advertising medium that was uncorrelated with sales). Difference-based heuristics were used less frequently but generated the most extreme allocations. Trend-based heuristics were used the least.

Keywords: managerial decision making, budget allocation, statistical graphics, marketing research, marketing metrics

Heuristics and Biases in Data-Based Decision Making: Effects of Experience, Training, and Graphical Data Displays

Managerial decision making is based on two fundamentally different types of information. The first type is the set of beliefs held by managers regarding what is generally true about the markets in which they compete. These beliefs stem from many sources, including formal training, everyday experience, and the high-level strategies formulated by company executives and industry analysts. Such "belief-based" information can be contrasted with the "data-based" information that routinely flows through most organizations (e.g., sales and profit figures, budgets and forecasts, various technical and market research results). Some of these data may conflict with managerial beliefs, leading to either an appropriate updating of the beliefs or an interpretation of the data biased in the direction of prior beliefs (e.g., Biyalogorsky, Boulding, and Staelin 2006; Bolton 2003). In addition to prior beliefs, decisions based on numerical data can be strongly biased by the cognitive heuristics used to analyze the data (e.g., Hutchinson and Alba 1997; Nisbett et al. 1983). These latter phenomena are the focus of this research.

The central role of numerical data in many managerial decisions prompts an important question about the extent to

*J. Wesley Hutchinson is Stephen J. Heyman Professor and Professor of Marketing, Wharton School, University of Pennsylvania (e-mail: jwhutch@wharton.upenn.edu). Joseph W. Alba is Distinguished Professor of Marketing, Warrington College of Business, University of Florida (e-mail: joe.alba@cba.ufl.edu). Eric M. Eisenstein is Assistant Professor of Marketing, Fox School of Business, Temple University (e-mail: eric.eisenstein@temple.edu). The authors thank the Wharton Behavioral Laboratory for support in data collection. They also thank Lauren Yourshaw, Gal Zauberman, and seminar audiences too numerous to list for their comments on previous versions of this article. Finally, the authors thank the anonymous *JMR* reviewers for their constructive comments and suggestions. John Lynch served as associate editor for this article.

ISSN: 0022-2437 (print), 1547-7193 (electronic)

which known biases can be reduced. Two obvious sources of bias reduction are explicit training and real-world experience (Arkes 1991; Boulding, Morgan, and Staelin 1997; Fong, Krantz, and Nisbett 1986; Kahn, Luce, and Nowlis 2006; Larrick 2004). In addition, many researchers have argued that data presentations using appropriate graphical formats can greatly enhance data comprehension, which presumably helps reduce bias (Lurie and Mason 2007; Wainer and Velleman 2001). The current research examines the ability of graphical presentation, experience, and training to reduce biases in data-based allocation decisions.

Building on the results of prior academic research and current trends in business analytics, our research goal is to answer three important questions about data-based inferences made by marketing managers. First, do biases exist that are sufficiently strong to lead to seriously suboptimal decisions? Second, do graphical data displays, real-world experience, or explicit training reduce such biases? Third, are the observed biases adequately explained by a relatively small set of natural heuristics that managers use when making data-based allocation decisions? From our results, we conclude yes, no, and yes, respectively.

LITERATURE REVIEW

In our experimental paradigm, the decision maker receives historical data about several resource variables (e.g., expenditures on several types of advertising) and an outcome variable (e.g., sales results) and must allocate a budget across the resource variables to maximize future outcomes. This task requires that the decision maker identify resource variables that are causally related to the decision outcomes—which in turn entails some assessment of the correlation between each resource variable and the outcome variable. Although people are often able to perceive large differences in correlations correctly (e.g., Broniarczyk and Alba 1994), they are also influenced by factors unrelated to the statistical properties of the data. In this section, we review three factors that are important for data-based inferences.[1]

Prior Beliefs

A large body of evidence suggests that prior beliefs significantly distort perceptions of correlation, such that people overestimate relationships that are expected and underestimate relationships that are unexpected (e.g., Broniarczyk and Alba 1994; Chapman and Chapman 1969). The damaging effects of prior beliefs are also manifest in managerial decision settings in the form of belief inertia, which can result in escalation bias (Biyalogorsky, Boulding, and Staelin 2006) and the failure of analytic thinking to overcome informal prior impressions (Bolton 2003). In the current research, we manipulate prior beliefs through the labels of the resource variables. This well-known source of bias provides a benchmark for assessing the size of biases due to other sources and the size of any reductions in bias due to graphical presentation, experience, or training.

Temporal Frame

The semantic frame of a data set is the person's knowledge about the phenomena being represented by numerical variables (Hutchinson and Alba 1997). For this research, we manipulate the semantic frame by portraying the variables as either cross-sectional or time series; therefore, we use the more specific term "temporal frame." Insofar as people attempt to assess correlation as an indicator of causation, allocation decisions should be unaffected by the meanings of the numbers in a data set, because these meanings do not affect the correlations.[2] However, Hutchinson and Alba (1997) find that a temporal frame can strongly influence allocation decisions through its influence on the heuristics used to assess correlation. In particular, they find that difference-based heuristics (e.g., comparing adjacent observations) are more common for time-series frames, but exemplar-based heuristics (e.g., mimicking the allocations of the observations with the best outcomes) are more common for cross-sectional frames. We discuss these heuristics in detail subsequently, but it is important to note that they are defined abstractly and focus primarily on which aspects of the stimulus data are used to make inferences about correlations. Thus, they apply to graphic formats as well as tables, enabling us to use a heuristic-based model of perceived correlation to explain the entirety of our results.

Graphical Formats

As statisticians frequently note, graphical formats may provide insights into data that are not easily derived from tabular representations (Smith et al. 2002; Wainer and Velleman 2001). And inasmuch as widely available spreadsheet software and statistical packages increasingly allow users to display data in graphical formats, the field of statistical graphics has become more prominent in many disciplines, including marketing (Lurie and Mason 2007). However, the universally acclaimed gold standard of data display is actually very old: It is Charles Joseph Minard's chart portraying the losses incurred by Napoleon's army in the Russian campaign of 1812 (e.g., Tufte 2001). Minard published his chart in 1869, and its accolades derive from the clarity with which a complex, multivariate data set can be apprehended (see Figure 1). The chart shows the size of the army as a function of geographic location, time, and temperature, highlighting the magnitude of the French losses with dramatic differences in line thickness for the same locations during the advance and retreat. It also allows those with a little background knowledge to infer the relative contributions of the Russian army's scorched earth tactics and the brutal Russian winter. Thus, this famous chart portrays the central question addressed in our research: How does the display of factual data affect causal inferences? The statistical graphics literature suggests that effective displays of data reduce miscomprehension and ambiguity, which in turn should reduce biased assessments of the correlations between variables.

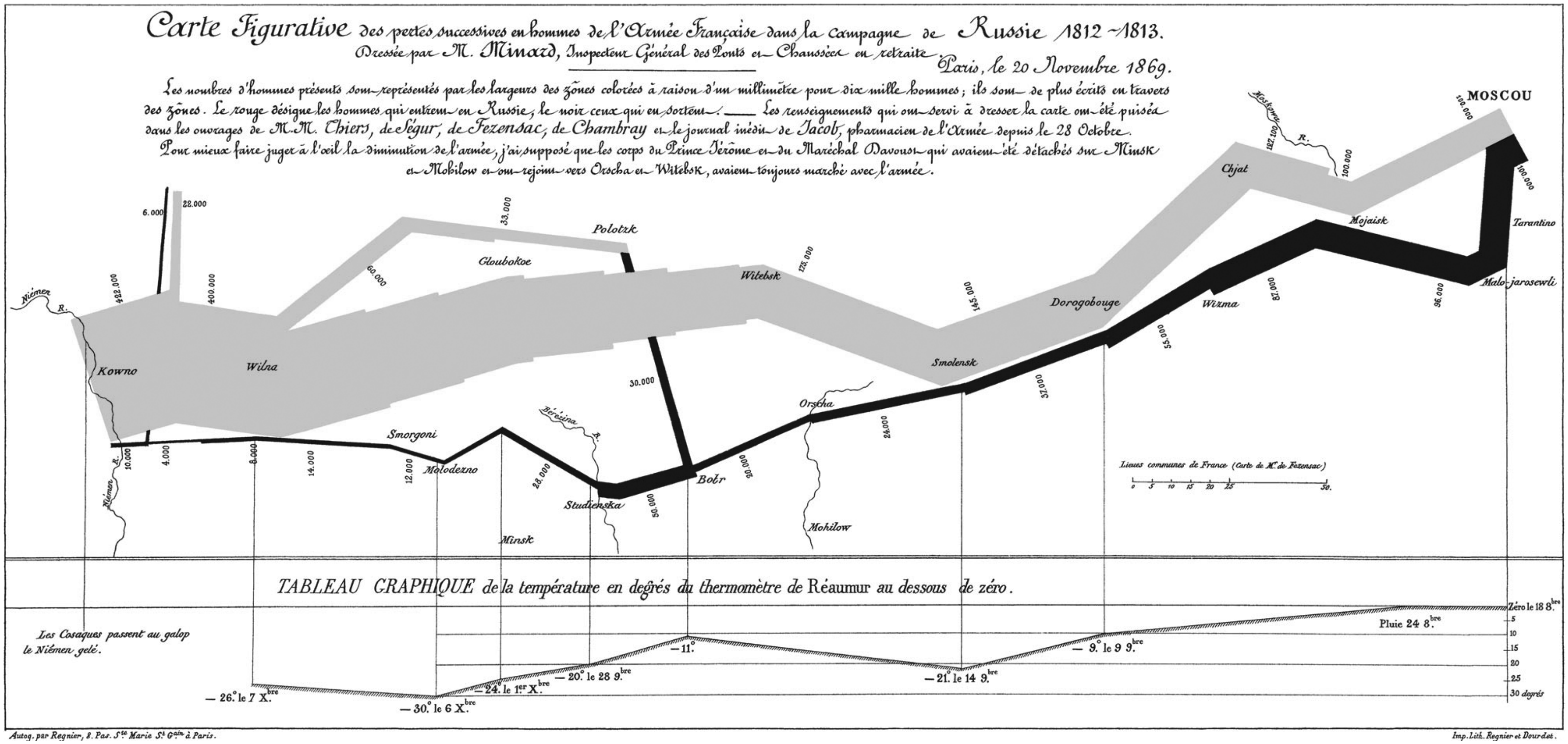
Perhaps because the virtues of graphical representation often appear self-evident, prior research has focused on the

[1]A more compete set of references is available in the Web Appendix (http://www.marketingpower.com/jmraug10).

[2]The choice of which statistical method or model specification is most appropriate often depends directly on the semantics of the data, but after a method or model is chosen, the meanings of the numbers do not influence the computations.

Figure 1

CHARLES JOSEPH MINARD'S CHART PORTRAYNG THE LOSSES SUFFERED BY NAPOLEON'S ARMY IN THE RUSSIAN CAMPAIGN OF 1812



Notes: The map's French caption reads: "Figurative Map of the successive losses in men of the French Army in the Russian campaign 1812–1813. Drawn up by M. Minard, Inspector General of Bridges and Roads in retirement. Paris, November 20, 1869. The numbers of men present are represented by the widths of the colored zones at a rate of one millimeter for every ten-thousand men; they are further written across the zones. The red [now gray] designates the men who enter into Russia, the black those who leave it. —— The information which has served to draw up the map has been extracted from the works of M. M. Thiers, of Segur, of Fezensac, of Chambray, and the unpublished diary of Jacob, pharmacist of the army since October 28th. In order to better judge with the eye the diminution of the army, I have assumed that the troops of prince Jerome and of Marshal Davoush who had been detached at Minsk and Moghilev and have rejoined around Orcha and Vitebsk, had always marched with the army." This graphic is freely available through Wikimedia Commons at http://en.wikipedia.org/wiki/File:Minard.png.

nuances of graphical processing rather than the influence of graphs in the larger decision environment. Statisticians offer useful advice regarding effective graphical portrayals of numerical information (e.g., Henry 1995; Tufte 2001; Wallgren et al. 1996) but are relatively silent with regard to the use of graphical formats to facilitate the comparison of different data patterns. Psychological research largely focuses on basic perceptual and comprehension processes of pattern recognition, mapping, and interpretation (e.g., Carpenter and Shah 1998; Simkin and Hastie 1987; Zacks and Tversky 1999). It also has focused on bivariate rather than multivariate data, rendering it less useful for the types of causal inference that we investigate here. Decision researchers similarly appreciate the potential importance of graphical formats but limit their investigations to predecisional processes, such as making point comparisons, identifying trends, and assessing risks (e.g., Chua, Yates, and Shah 2006; Harvey and Bolger 1996; Meyer, Shamo, and Gopher 1999). Other research that is cast in terms of budget allocation essentially examines perceptual issues (e.g., Benbasat and Dexter 1985).

In addition to conveying information more efficiently, graphical formats can alter the relative salience of different aspects of the data (Jarvenpaa 1990; Lane, Anderson, and Kellam 1985). These salience effects might facilitate accurate perceptions of the data and thereby dampen the biases created by prior beliefs and the temporal frame. Alternatively, graphical formats may make some aspects of the data more salient than others, leading to overemphasis and neglect of different decision variables. Given these opposing possible effects, a primary goal of our experiments is to test whether the graphic formats remediate or exacerbate the biases observed with tabular formats.

THE ALLOCATION TASK, OPTIMALITY, AND HEURISTIC-BASED BIASES

All the experiments asked participants to allocate an advertising budget across three media (newspaper, radio, and television). Allocations were relative (summing to 100) rather than absolute, both to simplify the task and because the profit implications of relative allocations are typically greater than those of budget size (Mantrala, Sinha, and Zoltners 1992). Data about previous allocations and the resulting sales outcomes were provided to support the decision. A typical example of the instructions follows:

> **Instructions:** Below are financial data for 16 department stores from the same chain in different cities. Data for each store include total store sales and the amounts of money spent on newspaper, radio, and television advertising. The 16 cities were approximately equal in size and sales potential, and there were no seasonal, inflationary or general economic fluctuations during this time. Therefore, management feels that store sales are a good indicator of advertising effectiveness. Examine these data to determine which advertising medium was most effective, then recommend a general policy for what percentage of the advertising budget should be spent advertising in each medium. ***<u>For this problem, be sure to allocate the largest percentage to the medium that was most effective and the smallest percentage to the medium that was least effective.</u>*** Record your allocation policy in the blanks provided below. Your three numbers should add up to 100.

The third and fourth sentences of the instructions were designed to emphasize a reliance on the presented data relative to prior beliefs. A key sentence, displayed to participants in bold, italicized, underlined font, emphasized causal thinking (i.e., effectiveness) and discouraged strategies that do not rely on data-based inferences. Moreover, it conveyed that the allocations were a "choice" regarding the most effective variable. This implicit choice provides a robust dependent measure, beyond the allocations themselves.

Stimulus Design and Optimal Allocations

The primary goal in constructing stimulus data for the allocation task was to create differences between optimal allocations and the allocations predicted by heuristic-based decision processes (see Table 1 and subsequent discussions). Figure 2 shows the three-step process used to create stimulus data. First, we linearly transformed the schematic design of the resource variables into realistic dollar values. Second, we computed the outcome variable as a linear function of the resource variables, plus a small amount of error (with no lagged effects).[3] Third, we represented the stimulus data in a chart or table (all stimuli are available in the Web Appendix at http://www.marketingpower.com/jmraug10).

Schematically, we refer to the three resource variables as the control variable, the global trend variable (which increased smoothly across observations), and the local contrast variable (which varied considerably from observation to observation without exhibiting a general trend). In all experiments, the control variable was uncorrelated with the outcome variable and the other two resource variables ($r \leq .1$), and the global trend and local contrast variables were equally correlated with the outcome variable but uncorrelated with each other. Thus, from a normative perspective, little or no money should be allocated to the control variable. Across stimulus designs, we varied the cost effectiveness of the global trend and local contrast variables using the linear transformations described previously, such that one or the other should be preferred from a normative perspective of providing more "bang for the buck."

The precise allocations that are optimal for a given data design depend on the specific assumptions made. For example, if the data were analyzed statistically, an analyst would find that a linear model predicts sales well. If linearity is assumed for the entire range of possible allocations, the strictly optimal policy is to allocate the entire budget to the resource variable with the largest regression coefficient (see the first row of Table 1). The strictly optimal policy is quite extreme; however, even more robustly optimal policies always favor the resource variable with the largest regression coefficient (see the second row of Table 1).[4] In particular, the global trend variable is favored for data designs 1 and 3, and the local contrast variable is favored for data

[3]Lagged, or carryover, effects are an important area of marketing strategy and empirical modeling. However, they greatly complicate the problem in terms of statistically appropriate analyses and optimal allocations, and therefore, we leave their investigation to further research.

[4]A more detailed discussion of robust policies is provided in the Web Appendix (http://www.marketingpower.com/jmraug10). Here, we note that both diffuse priors and concavity in the functions relating resource variables to the outcome variable militate toward allocating less than 100% to a single resource variable; however, allocations should be ordered consistently with the magnitudes of the regression coefficients.

Table 1

SUMMARY OF STIMULUS DESIGNS SHOWING HOW HEURISTIC-BASED ALLOCATIONS DIFFER FROM OPTIMAL ALLOCATIONS, TOGETHER WITH OBSERVED MEAN ALLOCATIONS AND PROPORTIONS OF PARTICIPANTS CHOOSING EACH RESOURCE VARIABLE AS MOST EFFECTIVE FOR EACH EXPERIMENT

Experiment	*1*			*2*			*3*		
Data Design	*1*			*2a and 2b, Averages*			*3*		
Resource Variable	*C*	*G*	*L*	*C*	*G*	*L*	*C*	*G*	*L*
Allocations Predicted by Normative Indexes									
Strictly optimal	0	**100**	0	0	0	**100**	0	**100**	0
Robustly optimal	2	**58**	40	2	40	**58**	2	**58**	40
Allocations Predicted by Heuristic-Based Indexes									
Difference-based	19	25	**56**	26	22	**53**	19	25	**56**
Trend-based	22	**45**	33	18	**48**	35	22	**45**	33
Exemplar-based	30	**40**	30	30	**43**	27	**48**	22	30
Observed Mean Allocations (Proportions of Participants Choosing Each Resource Variable as Most Effective)									
Grand Means	22	38	**40**	22	**42**	38	**37**	27	36
	(.05)	(**.48**)	(.37)	(.04)	(**.58**)	(.36)	(**.43**)	(.14)	(.31)
Temporal Frame									
Cross-sectional (EB)	20	39	**41**	23	**46**	32	**39**	26	35
	(.01)	(**.52**)	(.34)	(.03)	(**.77**)	(.16)	(**.48**)	(.12)	(.27)
Time series (DB and TB)	24	37	**39**	20	37	**42**	34	28	**37**
	(.08)	(**.43**)	(.39)	(.03)	(.40)	(**.53**)	(**.38**)	(.17)	(.34)
Graphic Format									
Tables	25	**40**	35	23	**42**	35	**37**	27	35
	(.05)	(**.54**)	(.27)	(.04)	(**.59**)	(.35)	(**.37**)	(.17)	(.31)
Bars (EB)	20	**43**	37	22	**48**	30	**39**	25	36
	(.03)	(**.60**)	(.30)	(.02)	(**.84**)	(.09)	(**.54**)	(.08)	(.28)
Lines (DB and TB)	20	30	**50**	18	34	**47**	33	29	**38**
	(.07)	(.30)	(**.54**)	(.02)	(.33)	(**.58**)	(**.46**)	(.15)	(.33)
Prior Beliefs									
Favor global	21	**40**	39	22	**45**	33	**38**	29	33
	(.02)	(**.54**)	(.33)	(.02)	(**.66**)	(.29)	(**.45**)	(.16)	(.28)
Favor local	22	36	**42**	21	**38**	41	36	25	**39**
	(.07)	(**.41**)	(.40)	(.04)	(**.50**)	(.40)	(**.41**)	(.13)	(.33)

Notes: Resource variables take the following labels: "C" for control, "G" for global trend, and "L" for local contrast. Predicted allocations, a_{Xi}, were computed from correlations, $r_{Xi,Y}$, of resource variables, X_i, with the outcome variable, Y, after the heuristic-based transformations of the stimulus data using a logit model, $a_{Xi} = \exp(r_{Xi},Y)/[\exp(r_{C,Y}) + \exp(r_{G,Y}) + \exp(r_{L,Y})]$, where a_{Xi} is the allocation for variable i = C, G, or L (except best exemplars heuristics, which use averages of observed allocations; for details, see the Web Appendix, http://www.marketingpower.com/jmraug10). For each participant, a variable was selected as most effective if it received the highest allocation (if two variables received the same highest value, neither was counted as having been selected). Bold values indicate the largest allocation in each condition. The DB, TB, and EB (difference-based, trend-based, and exemplar-based, respectively) variables indicate which heuristics were predicted to be more likely for each experimental factor (if any). Temporal Frame and Graphic Format were orthogonally varied in Experiment 1; in Experiments 2 and 3, all bar charts were cross-sectional, and all line charts were time series.

design 2. Actual allocations are biased when they differ from optimal allocations, especially when different resource variables are favored.

Heuristic-Based Allocations

It is widely accepted among social scientists that human decision making is "boundedly rational." Heuristic approximations of optimal decision processes are often used because the computations required for optimal decisions either cannot be justified from a cost–benefit perspective or are simply not within the realm of human ability. For budget allocation decisions based on tabular data, Hutchinson and Alba (1997) find three general types of frequently used heuristics: difference-based, trend-based, and exemplar-based. Consistent with findings on intuitive correlation assessments, all three heuristics enable people to distinguish large differences in correlations. This ability, plus some intuition about "bang for the buck," makes it plausible that people will be able to approximate optimal allocations. However, these heuristics also create specific biases that might lead people astray. In this section, we describe each type of heuristic and the pattern of allocations they predict for our stimulus data (see Table 2 for a simple example and the Web Appendix, at http://www.marketingpower.com/jmraug10, for details).[5]

Difference-based heuristics examine local changes in allocations for each resource variable and compare those changes with associated changes in the outcome variable. For the data in Figure 2, the direction of changes in the local contrast variable matches the direction of changes in the outcome variable for 14 of 15 adjacent differences, compared with 2 of 15 for the global trend variable and 8 of 15 for the control. Thus, difference-based heuristics favor the local contrast variable. This bias favoring the local contrast variable occurs for all our data designs (see Table 1).

[5]Hutchinson and Alba (1997) fit 36 versions of six types of heuristics to their data in determining that these three were best fitting across their experiments. They also model heterogeneity in two ways and conclude that at least some people used each of these three types in most conditions.

Figure 2
ILLUSTRATION OF THE STIMULUS CONSTRUCTION PROCESS

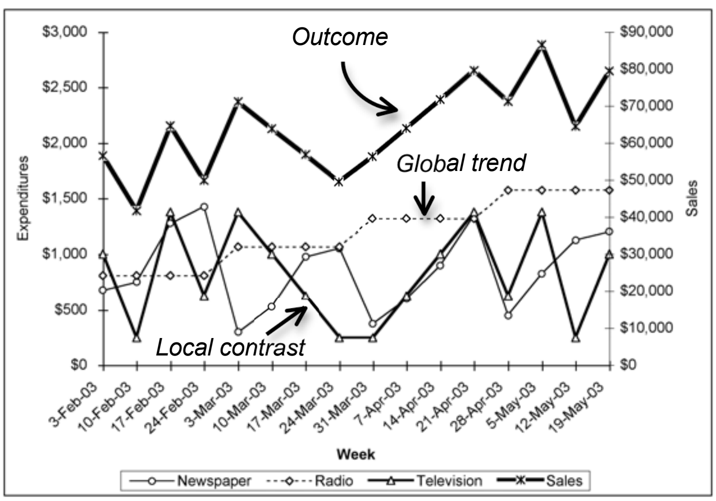
Schematic Design for Resource Variables

Control	Global Trend	Local Contrast
6	1	3
7	1	1
14	1	4
16	1	2
1	2	4
4	2	3
10	2	2
11	2	1
2	3	1
5	3	2
9	3	3
15	3	4
3	4	2
8	4	4
12	4	1
13	4	3

Stimulus Data Created Through Linear Transformations of the Schematic Design

Outcome	Control	Global Trend	Local Contrast
$56,535	$675	$806	$1,000
$41,692	$750	$806	$250
$64,594	$1,275	$806	$1,375
$49,826	$1,425	$806	$625
$71,078	$300	$1,063	$1,375
$63,844	$525	$1,063	$1,000
$56,835	$975	$1,063	$625
$49,451	$1,050	$1,063	$250
$56,235	$375	$1,319	$250
$63,919	$600	$1,319	$625
$71,678	$900	$1,319	$1,000
$79,588	$1,350	$1,319	$1,375
$71,228	$450	$1,575	$625
$86,522	$825	$1,575	$1,375
$64,444	$1,125	$1,575	$250
$79,438	$1,200	$1,575	$1,000

Final Stimuli Constructed by Adding Semantic Frame, Format, and Prior Belief



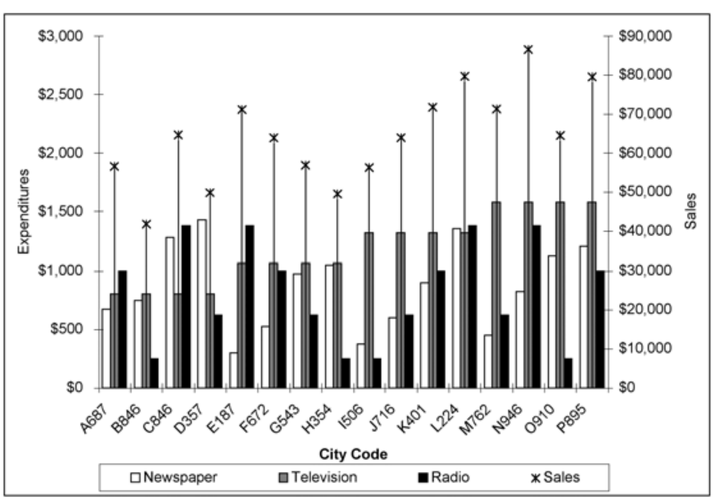


Table 2

A SIMPLE EXAMPLE OF ALLOCATION HEURISTICS

Variable	*Observations* a	b	c	d	e	*Correlation with Outcomes*	*Predicted Allocations*
Global trend (G)	1	2	3	4	5	.72	
Local contrast (L)	1	4	2	4	1	.69	
Outcome (G + L)	2	6	5	8	6		
Difference-Based Heuristic (Between Adjacent Observations)							
Global trend	—	1	1	1	1	.00	.27
Local contrast	—	3	–2	2	–3	1.00	.73
Outcome	—	4	–1	3	–2		
Trend-Based Heuristic (Smoothing by Averaging Adjacent Observations)							
Global trend	—	1.5	2.5	3.5	4.5	.98	.68
Local contrast	—	2.5	3	3	2.5	.22	.32
Outcome	—	4	5.5	6.5	7		
Exemplar-Based Heuristic (Average Allocations for the Best Three Observations)						*Average Allocation*	
Global trend	—	2	—	4	5	3.7	.55
Local contrast	—	4	—	4	1	3.0	.45
Outcome	—	6	—	8	6		

Notes: For difference-based and trend-based heuristics, predicted allocations, a_{Xi}, were computed from correlations, $r_{Xi, Y}$, of resource variables, X_i, with the outcome variable, Y, after the heuristic-based transformations of the stimulus data using a standard multinomial logit model, $a_{Xi} = \exp(r_{Xi, Y})/[\exp(r_{G, Y}) + \exp(r_{L, Y})]$, where a_{Xi} is the allocation for variable i = G or L. For the best exemplars heuristic, predicted allocations were the average of the observed allocations for the three observations with the highest values for the outcome variable (for more details, see Figure A1 and the discussion in the Web Appendix, http://www.marketingpower.com/jmraug10).

Hutchinson and Alba (1997) find that difference-based heuristics are especially likely for time-series data. Indeed, line charts are frequently recommended for time-series data precisely because they highlight changes from one period to another (e.g., Henry 1995; Wallgren et al. 1996). Thus, line charts should increase the use of difference-based heuristics.

Rather than examining local changes, people might "smooth" the data and apprehend general trends (e.g., increasing, U shaped). In all four data designs, both the outcome variable and the global trend variable exhibit a generally increasing trend that is not present in the other two variables. Thus, trend-based heuristics always favor the global trend variable (see Table 1). Hutchinson and Alba (1997) find evidence of this heuristic, but the frequency of use is far less than for difference-based or exemplar-based heuristics.[6] In line with the statistical graphics literature (e.g., Henry 1995), we expect that line charts will increase the use of trend-based heuristics.

Finally, Hutchinson and Alba (1997) note a rather different—but naturally appealing—heuristic, the best exemplars heuristic, that is also commonly used (especially for cross-sectional data). This heuristic does not assess correlation per se but rather attempts, in the spirit of best practices benchmarking, to imitate success. The best exemplars heuristic computes the average allocations for the observations with the highest values on the outcome variable and uses these averages as the basis for future allocations. Exemplar-based heuristics have been found to explain a variety of behaviors (e.g., Juslin, Olsson, and Olsson 2003; Meyer 1989; Van Osselaer, Janiszewski, and Cunha 2004). For our stimuli, exemplar-based heuristics favor the global trend variable in Experiments 1 and 2 but favor the control variable in Experiment 3 (see Table 1).

Summary

The top portion of Table 1 shows the predicted allocations for normative and heuristic-based indexes across the three data designs in our experiments. None of the three types of heuristics agree with optimal allocations across all designs. *Optimal allocations* favor the global trend variable for designs 1 and 3 and the local contrast variable for design 2. *Difference-based heuristics* always favor the local contrast variable; *trend-based heuristics* always favor the global trend variable; and *exemplar-based heuristics* favor the global trend variable for designs 1 and 2 but the control variable for design 3. Thus, each allocation process has a unique pattern of predictions across our experiments, which enables clear assessment of the use of different heuristics. We expect that difference-based and trend-based heuristics will be prominent in time-series frames and line chart formats but that exemplar-based heuristics will be prominent in cross-sectional frames and bar chart and tabular formats. None of these predictions are consistent with optimal allocations.

EXPERIMENT 1: BIASED ALLOCATIONS AND THE EFFECTS OF GRAPHIC FORMAT AND REAL-WORLD EXPERIENCE

Experiment 1 had two primary goals. The first goal was to assess the influence of the graphical format on data-based inferences, benchmarked against the documented effects of the temporal frame and prior beliefs. Of particular interest was determining whether graphical formats reduced biases previously observed with tables. The second goal was to determine whether real-world experience reduced any

[6]To create the global trend variable, we sorted observations by that variable. This step necessarily reduces local variation in the sort variable and induces matched local variation between the outcome and local contrast variables (for a detailed discussion of this effect of sorting, see Diehl, Kornish, and Lynch 2003). We have replicated the results reported subsequently for trend variables not created by sorting (see Experiment A1 in the Web Appendix at http://www.marketingpower.com/jmraug10).

observed biases. Decision biases can sometimes be reduced by high levels of domain-specific knowledge, obtained from either formal training or real-world experience (e.g., Arkes 1991; Fong, Krantz, and Nisbett 1986). Alternatively, knowledge increases some types of decision biases, particularly those due to overconfidence (e.g., Alba and Hutchinson 2000). In Experiment 1, we compared the performance of students with that of marketing managers.

Method

Participants. A total of 157 students from the University of Pennsylvania received $10 compensation for participating in the experiment (which included other studies not related to the current research). In addition, 500 randomly selected, nonacademic members of the American Marketing Association were mailed invitations to participate in an online study. The incentive for the managerial sample was a $20 donation, to be made to a charitable organization of the participant's choice, and a chance to win free admittance to a one-week executive education program. Eighty-nine managers participated. The student and managerial samples participated through identical Web sites.

Instructions, procedure, and stimuli. Only the first sentence of the instructions varied across conditions. For the cross-sectional conditions, each observation was described as financial data from 1 of 16 department stores from the same chain. For the time-series conditions, each observation was 1 of 16 weeks of financial data for a single department store. The remainder of the instructions were essentially identical to the example provided previously. We manipulated prior beliefs by switching the labels of the global trend and the local contrast variables. Preliminary tests using a different sample revealed that most participants expected television to be more effective than radio; thus, the global trend variable was favored when it was labeled "Television" (and the local contrast variable was labeled "Radio"; see Figure 2), whereas the local contrast variable was favored when the labels were reversed. All observations were presented simultaneously in a chart or table (see the Web Appendix at http://www.marketingpower.com/jmraug10). After participants made their budget allocations, a final page requested them to "describe how you determined which advertising medium was most effective and how you made your final budget allocation." A text box was provided for this response.

Experimental design. Participants were randomly assigned to 1 of 12 conditions that resulted from crossing three factors: graphical format (bars, lines, table), temporal frame (cross-sectional, time series), and prior beliefs (global trend variable favored, local contrast variable favored). Experience (i.e., student versus manager) was the fourth design factor.

Results

In Table 1, we provide a summary of the results for two dependent measures: allocations made to each resource variable and the proportions of participants choosing each variable as the most effective. The choice proportions do not sum to 1 because ties (i.e., two variables receiving the same maximum allocation) were not assigned to any variable. As we noted previously, difference-based heuristics favored the local contrast variable, but trend-based and exemplar-based heuristics favored the global trend variable. Therefore, in addition to analyses of allocations to each variable, we assessed the differences between allocations (i.e., global trend minus local contrast) using a between-subjects analysis of variance (ANOVA). Because few people chose the control as most effective, we only report binary logistic regression analyses for the choice of the global trend or the local contrast variables as most effective. All analyses used the fully crossed four-factor model with standard effects coding, as implemented by the GLM and CATMOD procedures in SAS.

Overall, the results indicated that suboptimal allocations were frequent. For the data design used in Experiment 1, allocations should be highest for the global trend variable to be considered optimal. However, only 48% of participants chose the global trend variable as most effective, and the mean allocation to the global trend variable was slightly less than the mean allocation to the local contrast variable (see Table 1). The effects of prior beliefs and the temporal frame were as predicted but small compared with the effects of graphic format. More important, graphic format did not reduce bias and, in some respects, increased it.

Prior beliefs. The effects of prior beliefs were as expected but weak. The mean allocation for the variable labeled "Television" was 41, and it was chosen as most effective by 47% of participants, whereas the mean allocation for the variable labeled "Radio" was 37, and it was chosen as most effective by 36% of participants. The effect of prior beliefs for allocation differences only approached statistical significance ($F(1, 222) = 2.3$, $p = .13$, mean standard error [MSE] = 1064) and was not significant for either choice measure.

Temporal frame. Similar to the results for prior beliefs, the effect of the temporal frame was as expected but weak. The difference in allocations was not significant; however, the global trend variable was chosen as most effective more often for cross-sectional than for time-series data (52% versus 43%, $\chi^2_1 = 3.7$, $p = .05$), and the local contrast variable was chosen as most effective more often for time-series than for cross-sectional data (39% versus 34%, $\chi^2_1 = 1.3$, $p = .25$).

Graphic format. In contrast with the effects of prior beliefs and temporal frame, the effect of graphical format was very strong. Mean allocations to the global trend variable were 40 for tables and 43 for bar charts but only 30 for line charts ($F(1, 222) = 28.7$, $p < .0001$, MSE = 351 for bars and tables versus lines). Conversely, mean allocations to the local contrast variable were 35 for tables and 37 for bar charts but increased to 50 for line charts ($F(1, 222) = 27.7$, $p < .0001$, MSE = 251 for bars and tables versus lines). Similarly, the global trend variable was chosen as most effective more often for tables and bar charts than for line charts (54%, 60%, and 30%, respectively; $\chi^2_1 = 13.9$, $p = .0002$ for bars and tables versus lines), and the local contrast variable was chosen as most effective less often for tables and bar charts than for line charts (27%, 30%, and 54%, respectively; $\chi^2_1 = 12.9$, $p = .0003$ for bars and tables versus lines). Thus, there was no evidence that graphical formats reduced bias. Indeed, line charts increased bias in two senses. First, optimal allocations favored the global trend variable, but line charts resulted in a strong preference for the local contrast variable. Second, allocations were more extreme for line charts. The mean difference in allocations to global trend and local contrast variables was –20 for

line charts compared with +6 for tables and +7 for bar charts ($F(1, 222) = 31.8$, $p < .0001$, MSE = 1064 for lines versus tables and bars).

Allocations based on line charts were even more extreme at the individual level. Figure 3 shows the marginal distributions of the differences in allocations to the global trend and local contrast variables for each type of graphic format.[7] The distribution for line charts had a large mode around –60, in addition to a mode around +10. The distribution for tables was more or less symmetric around a mode at +10. The distribution for bar charts was slightly skewed, with a mode near +20. Thus, bar charts appeared to have helped reduce bias compared with line charts and tables because the optimal allocations favored the global trend variable. However, we test this hypothesis directly in Experiment 2 and find that bar charts continue to favor the global trend variable, even when the optimal allocations favor the local contrast variable. Consequently, we are more inclined to interpret the increased extremity of allocations for bar charts compared with tables as increased bias (similar to that of line charts, but in the opposite direction), rather than as indicative of greater insight regarding optimal allocations. Overall the graphic formats hurt more than they helped.

Real-world experience. Managers took longer to make their decisions than students (212 seconds versus 130 seconds; $F(1, 219) = 22.5$, $p < .0001$, MSE = 15,955). However, similar to the results for graphic format, there was little evidence that real-world experience reduced bias. There were no main effects of experience for allocations or for the variable chosen as most effective. However, follow-up analyses revealed significant effects when the data were presented in tables, not in charts, as Figure 4 shows. When the data were presented in tables, students strongly favored the global trend variable, whereas managers slightly favored the local contrast variable.[8] Thus, if anything, managers were more suboptimal than students because optimal allocation favors the global trend variable. However, as we discussed in the context of the effects of bar and line charts, the extremity of the allocation was another indicator of suboptimality. Managers were less extreme in their allocations than students when the data were presented in tables, which we interpreted as a beneficial difference. However, this small advantage disappeared when the data were graphically displayed, again indicating that graphic formats hurt more than they helped.

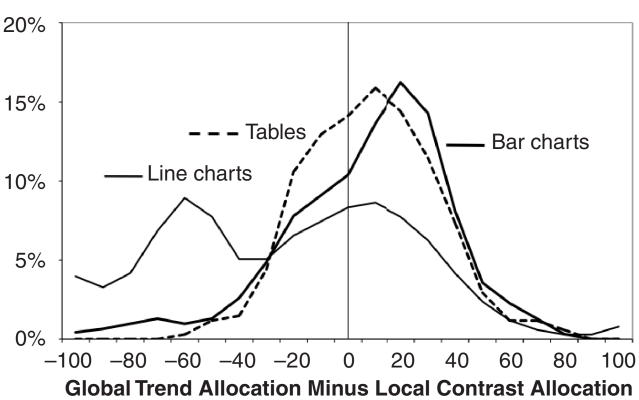
Open-ended responses. The open-ended responses were coded according to 13 categories (see Table 3). Nine categories represented allocation heuristics. In addition to statements that clearly described (1) adjacent differences, (2) trend-based heuristics, and (3) best exemplars heuristics, participants explained their allocations with reference to (4) correlations, (5) differences that were not clearly described as adjacent, (6) best and worst exemplars, (7) reasoning that was conditional on the independent variables (e.g., "I looked at the weeks with the highest spending on each type of advertising"), (8) cost effectiveness (e.g., "bang for the buck"), (9) guessing, and (10) heuristics that were not classifiable. Another three categories tracked (11) mentions of prior knowledge, (12) mentions of the total level of adver-

[7]We provide a more complete analysis of the extremity of allocations apparent in Figure 3 in the "General Discussion" section; for a more general discussion of unobserved heterogeneity in behavioral research, see Hutchinson, Kamakura, and Lynch (2000).

[8]This effect of experience when the data were presented in tables was significant for allocation differences (12 for students, 0 for managers; $F(1, 77) = 6.3$, $p = .01$, MSE = 450) and for the proportion choosing the global trend variable as most effective (.64 for students, .37 for managers; $\chi^2_1 = 5.7$, $p = .02$) and was marginally significant for the proportion choosing the local contrast variable as most effective (.21 for students, .40 for managers; $\chi^2_1 = 3.5$, $p = .06$).

Figure 3

HETEROGENEITY IN THE EXTENT TO WHICH THE GLOBAL TREND VARIABLE WAS FAVORED OVER THE LOCAL CONTRAST VARIABLE IN EXPERIMENT 1, AS INDICATED BY THE MARGINAL DISTRIBUTIONS OF THE DIFFERENCE BETWEEN GLOBAL TREND AND LOCAL CONTRAST ALLOCATIONS



Notes: Curves represent means within each ten-point bin, smoothed by computing weighted averages with adjacent bins, using weight ratios of 1:2:1.

Figure 4

EFFECTS OF GRAPHIC FORMAT AND REAL-WORLD EXPERIENCE ON ALLOCATIONS IN EXPERIMENT 1

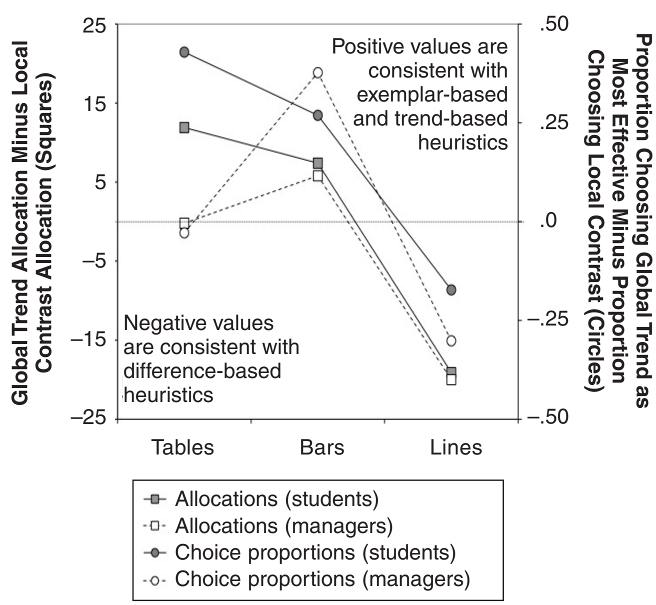


Table 3
SUMMARY OF OPEN-ENDED RESPONSES FOR EXPERIMENT 1

		Response Code Frequency as a Function of Experimental Design Variables[a]									*Regression Coefficients for Allocation Differences as a Function of Response Codes*[b]
		Frame		*Format*			*Prior Belief*		*Experience*		
	Response Code Frequency (% of Subjects)	*Cross-Sectional*	*Time Series*	*Bars*	*Lines*	*Table*	*Favors Local*	*Favors Global*	*Students*	*Managers*	
Correlation	.17	.16	.19	.14	.29	.08**	.21	.13†	.13	.25*	–14**
Difference-Based Heuristics											
Adjacent differences	.12	.06	.19**	.08	.24	.05***	.14	.10	.11	.15	–37***
Nonspecific differences	.13	.07	.19**	.08	.19	.11†	.13	.12	.10	.17†	–15*
Trend-Based Heuristics											
Trend	.06	.02	.09*	.03	.13	.01**	.03	.08†	.05	.08	11
Exemplar-Based Heuristics											
Best exemplars	.15	.23	.06***	.21	.02	.21**	.13	.16	.17	.10†	18**
Best/worst exemplars	.06	.10	.02*	.05	.01	.11†	.05	.07	.06	.06	13
Other Heuristics											
Conditional reasoning	.11	.11	.10	.18	.01	.13*	.12	.09	.12	.08	9
Bang for the buck	.06	.07	.04	.05	.07	.05	.04	.08	.03	.10*	–4
Guessed	.02	.01	.02	.01	.04	.00	.02	.02	.01	.02	37**
Cannot classify	.32	.29	.35	.29	.32	.35	.34	.30	.38	.21**	21***
Other Responses											
Prior knowledge	.11	.11	.12	.10	.10	.14	.10	.13	.06	.21***	6
Total spending	.02	.03	.02	.01	.01	.05	.02	.03	.02	.02	4
Data/analysis complaint	.04	.02	.06	.04	.06	.02	.04	.04	.01	.10*	18*

†$p < .25$.
*$p < .05$.
**$p < .01$.
***$p < .001$.
[a]Superscripts on the final value of a given effect indicate statistical significance of the chi-square test for that effect.
[b]Allocation differences were computed as allocations to the global trend variable minus allocations to the local contrast variable. The intercept in the model was –7. Superscripts on each coefficient indicate statistical significance of a t-test against 0.

tising expenditures, and (13) complaints about the data or the inability to conduct a statistical analysis.[9]

We conducted two types of analyses of these responses. First, we modeled each code separately as a function of the experimental factors and level of expertise. The results of these analyses were consistent with our prior discussion (see Table 3). Difference-based and trend-based heuristics were more frequently mentioned for line charts and time-series data. Exemplar-based heuristics were more frequently mentioned for bar charts and tables and cross-sectional data. Real-world experience was associated with a reliance on prior beliefs, as we expected. Experience also increased the tendencies to use correlation, consider cost effectiveness, be more articulate (i.e., produce classifiable responses), be concerned about the dangers of "eye-balling" data, and avoid the use of a best exemplars heuristic. All these effects were normatively positive but evidently did not translate into better allocation decisions (with the possible exception of when data were presented in tables).

Second, we regressed allocation differences (i.e., allocation to global trend minus allocation to local contrast for each participant) onto dummy variables for the response codes (see the rightmost column of Table 3). The results were consistent with our assumptions about allocation heuristics. Adjacent-differences and nonspecific-differences heuristics had negative coefficients, whereas trend-based and exemplar-based heuristics had positive coefficients. Thus, participants' allocations were generally consistent with their verbalized decision heuristics.

Trend neglect. The analyses of the open-ended responses suggest that difference-based and exemplar-based heuristics were relatively frequent, but trend-based heuristics were rare. First, trend-based heuristics were mentioned considerably less frequently than difference-based and exemplar-based heuristics (6%, 25%, and 21%, respectively). Second, in the regression analysis of the allocation differences, the coefficient for trend-based heuristics was not significant. Hutchinson and Alba (1997) report similar results. We refer to this phenomenon as trend neglect and investigate it further in Experiments 2 and 3. Regarding the effects of graphic formats, our results support the somewhat ironic conclusion that line charts had the potential to remediate bias because they facilitated both difference-based and trend-based heuristics (as predicted by statistical graphics literature); however, difference-based heuristics were so much stronger and more frequently used than trend-based heuristics that the net result was a stronger bias than the one caused by exemplar-based heuristics.

Explicit training. The results of Experiment 1 are somewhat disconcerting, insofar as graphic formats exerted much stronger effects than those of prior beliefs or temporal frame, but they increased, rather than decreased, the extremity of allocations and, for line charts, made them far from optimal. In addition, real-world experience did not reduce this bias. These results are similar to those of Boulding, Morgan, and Staelin (1997), who find that enriching and

[9]Blinded to experimental condition and level of expertise (except as occasionally revealed by the response itself), one of the authors initially coded the responses. Then, another author (also blinded to condition and expertise) reviewed all those classifications. Agreement was 82% (68% for an independently coded subsample and 87% when the second coder explicitly checked the first coder). Discrepancies were resolved by discussion.

improving the information environment had minimal effects on escalation-of-commitment biases. These authors also found that interventions that affected the rule used for decision making were more successful. Therefore, we conducted a follow-up experiment that used explicit training in an attempt to reduce bias (see the Web Appendix at http://www.marketingpower.com/jmraug10).

The training defined the three heuristics, warned against best exemplars, recommended the combined use of difference-based and trend-based heuristics, and provided participants with specific examples of different levels of correlation, in addition to verbal instructions about the heuristics. A control version of training simply exhorted participants to examine all the data and think carefully. The allocation task instructions were similar to those of Experiment 1. A performance-related monetary incentive also was included. Despite all these efforts, we again observed biases similar to those of Experiment 1, and explicit training had no beneficial effects. Similarly, covariates related to academic training (e.g., number of course in statistics or business) were not correlated with the degree of bias.

EXPERIMENT 2: CONFIRMING HEURISTIC-BASED SUBOPTIMALITY

In Experiment 1, line charts increased the use of difference-based heuristics, and the end result was a suboptimal allocation because difference-based heuristics favored the local contrast variable (in contrast with the optimal allocation, which favored the global trend variable). In Experiment 2, we adjusted the linear transformations used to create the data design so that the local contrast variable, rather than the global trend variable, was optimal. Specifically, in Experiment 1 the sales regression coefficients for the global trend and local contrast variables were 28 and 20, respectively (making the global trend variable more cost effective). In Experiment 2, the coefficients were 20 and 28, respectively (making the local contrast variable more cost effective). Experiment 2 also provided an opportunity to replicate the findings of Experiment 1 using different data designs.

Method

Participants and design. A total of 213 undergraduate volunteers who were taking an introductory marketing course at the University of Pennsylvania participated in Experiment 2. The experiment used a $2 \times 2 \times 2 \times 2$ between-subjects factorial design. The factors were prior beliefs (favored the global trend variable, favored the local contrast variable), temporal frame (cross-sectional, time series), graphical format (chart, table), and data design (2a, 2b).[10] To simplify the design but still create strong effects, and in line with the results of Experiment 1, the cross-sectional chart was a bar chart, and the time-series chart was a line chart. This mapping is consistent with current recommendations in statistical graphics literature.

Instructions and stimuli. Other than the changes in the schematic design and linear transformations, the stimuli were the same as in Experiment 1. The instructions also were the same as in Experiment 1, except that Experiment 2 used a paper-and-pencil format, and no open-ended responses were collected.

Results

The most important result emerged from the grand means (see Table 1). In Experiment 1, the grand means were generally consistent with optimal allocations: 48% of participants chose the global trend variable as most effective, compared with 37% who chose the local contrast variable. In Experiment 2, optimal allocations favored the local contrast variable, but the observed grand means again favored the global trend variable. The global trend variable was chosen as most effective more often (58% versus 48%, $\chi^2_1 = 5.4$, $p = .02$, according to an analysis that pooled data from both experiments). Thus, we ruled normatively rational budget allocations out as explanations for the observed grand means in Experiments 1 and 2, whereas heuristic-based explanations received strong support.

Replicating Experiment 1: prior beliefs. The effects of prior beliefs were as expected; however, they were larger and more reliable than in Experiment 1. Allocations to each variable were greater when labeled “Television” than “Radio” (45 versus 38 for global trend and 41 versus 33 for local contrast; $F(1, 197) = 11.9$, $p < .0007$, MSE = 204, and $F(1, 197) = 15.0$, $p < .0001$, MSE = 236, respectively). Similarly, both the global trend and the local contrast variables were chosen as most effective more often when labeled “Television” than “Radio” (66% versus 51%; $\chi^2_1 = 6.8$, $p = .009$, and 40% versus 29%; $\chi^2_1 = 3.5$, $p = .06$, respectively).

Replicating Experiment 1: temporal frame and graphic format. We expected the effect of the temporal frame to be much greater for charts than for tables because temporal frame was deliberately confounded with type of chart (i.e., bar charts for cross-sectional data and line charts for time-series data). We observed this interaction for both allocations and choice proportions ($F(1, 197) = 8.6$, $p = .004$, MSE = 777 for allocation differences; $\chi^2_1 = 4.4$, $p = .03$ for the choice of the global trend variable as most effective; and $\chi^2_1 = 3.6$, $p = .06$ for the choice of the local contrast variable).

The effect of the temporal frame was greater for charts than for tables for both allocations to the global trend variable (48 versus 34 for cross-sectional versus time series for charts; $F(1, 197) = 23.1$, $p < .0001$, MSE = 204, and 44 versus 40 for tables; $F(1, 197) = 2.4$, $p = .12$, MSE = 204) and allocations to the local contrast variable (30 versus 47 for cross-sectional versus time series for charts; $F(1, 197) = 30.7$, $p < .0001$, MSE = 236, and 33 versus 38 for tables; $F(1, 197) = 2.3$, $p = .13$, MSE = 236). Similarly, the effect of the temporal frame was greater for charts than for tables for both the choice of the global trend variable as most effective (84% versus 33% for cross-sectional versus time series for charts; $\chi^2_1 = 22.3$, $p < .0001$, and 70% versus 47% for tables; $\chi^2_1 = 6.5$, $p = .01$) and the choice of the local contrast variable as most effective (9% versus 58% for cross-sectional versus time series for charts; $\chi^2_1 = 20.0$, $p < .0001$, and 23% versus 48% for tables; $\chi^2_1 = 7.8$, $p = .005$). In

[10]The schematic values for data design 2a were the same as those for data design 1; the schematic values for data design 2b differed such that the allocations to the control variable distinguished between trend-based and exemplar-based heuristics. The two designs were essentially the same for allocations to the other two variables. The results of this manipulation supported trend neglect, but Experiment 3 provides more definitive results, so data designs 2a and 2b and the results for allocations to the control variable appear only in the Web Appendix (http://www.marketingpower.com/jmraug10).

absolute terms, the combined effects of temporal frame and graphic format were large for both dependent measures. From a normative perspective, neither temporal frame nor graphic format should affect allocations. Thus, the presence of heuristic-based biases is strongly indicated.

Summary. The main contribution of Experiment 2 was to rule out optimal allocation as an explanation for the observed pattern of allocations. Optimal allocation favored the local contrast variable in Experiment 2 (rather than the global trend variable favored in Experiment 1). However, the global trend variable was chosen as most effective even more often than in Experiment 1. Experiment 2 also replicated (with greater size and reliability) the effects of the format, frame, and prior beliefs that we observed in Experiment 1.

EXPERIMENT 3: MAXIMIZING THE EFFECTS OF TREND NEGLECT

In Experiments 1 and 2, both trend-based and exemplar-based heuristics predicted that the global trend variable would receive the largest allocation. We designed Experiment 3 such that the use of exemplar-based heuristics would lead to the choice of the control variable as most effective, trend-based heuristics would continue to favor the global trend variable, and difference-based heuristics would continue to favor the local contrast variable. To create the stimulus data for this experiment, we altered the linear transformations used for data design 1 by adding a constant value to the control variable and subtracting a constant value from the global trend and local contrast variables. Table 1 shows the resulting predicted allocations for each type of heuristic. It also illustrates the great weakness of all exemplar-based heuristics: They are heavily influenced by the overall means of each resource variable, regardless of the statistical relationship between the variables and outcomes. For data design 3, with exemplar-based heuristics, the result should be large allocations to the control variable, which is uncorrelated with sales. Such a result is extremely suboptimal.

Method

Participants and design. A total of 160 students from the University of Pennsylvania received $10 compensation for participating in Experiment 3 (along with several other studies not related to the current research). The experiment used a 2 × 2 × 2 between-subjects factorial design. The factors were prior beliefs (favored the global trend variable, favored the local contrast variable), temporal frame (cross-sectional, time series), and graphical format (chart, table). As in Experiment 2, the cross-sectional chart was a bar chart, and the time-series chart was a line chart.

Instructions and stimuli. Other than the changes in the linear transformations we described previously, the stimuli were the same as in Experiment 1. Instructions and the online procedure also were the same as in Experiment 1.

Results

As in Experiment 2, the most important results pertained to the grand means (see Table 1).[11] Consistent with the frequent use of exemplar-based heuristics, the mean allocation to the control variable increased from an average of 22 in Experiments 1 and 2 to an average of 37 in Experiment 3. To examine this increase statistically, we pooled the data across all three experiments. An ANOVA of allocations to the control variable revealed a significant effect of experiment ($F(2, 595) = 75.7$, $p < .0001$, MSE = 167), due almost entirely to the difference between Experiment 3 and the first two experiments ($F(1, 595) = 151.3$, $p < .0001$, MSE = 167). Similarly, the proportion of participants who chose the control variable as most effective increased from an average of .04 in Experiments 1 and 2 to .43 in Experiment 3. This tenfold increase was statistically significant ($\chi^2_2 = 79.8$, $p < .0001$) and again was due almost entirely to the difference between Experiment 3 and the first two experiments ($\chi^2_1 = 79.7$, $p < .0001$).

Conversely, the results were consistent with very infrequent use of trend-based heuristics. The mean allocation to the global trend variable decreased from an average of 40 in Experiments 1 and 2 to an average of 27 in Experiment 3. An ANOVA of allocations to the global trend variable revealed an effect of experiment ($F(2, 595) = 41.7$, $p < .0001$, MSE = 243), due almost entirely to the difference between Experiment 3 and the first two experiments ($F(1, 595) = 81.2$, $p < .0001$, MSE = 243). Similarly, the proportion of participants who chose the global trend variable as most effective decreased from an average of .54 in Experiments 1 and 2 to .14 in Experiment 3. This fourfold decrease was statistically significant ($\chi^2_2 = 65.0$, $p < .0001$) and again was due almost entirely to the difference between Experiment 3 and the first two experiments ($\chi^2_1 = 62.1$, $p < .0001$). Thus, trend neglect was extreme in Experiment 3. Moreover, these results, together with the open-ended responses from Experiment 1, suggest that comparably extreme trend neglect characterized the first two experiments but was hidden in the allocation results because both trend-based and exemplar-based heuristics favored the global trend variable (also see the analysis of the control variable allocations for Experiment 2 in the Web Appendix at http://www.marketingpower.com/jmraug10).

GENERAL DISCUSSION

The experimental findings we report here provide strong support for several conclusions about the ways managers use numerical data in budget allocation decisions. First, the allocations were biased compared with the benchmark of optimal allocations. Second, the biases were caused by the heuristics that decision makers used. Third, of the three types of heuristics identified in our data (i.e., difference-based, trend-based, and exemplar-based), difference-based and exemplar-based heuristics were used much more frequently than trend-based heuristics. Fourth, graphical formats that followed existing recommendations for the appropriate display of data did not reduce heuristic-dependent biases compared with data presented in tables. Line charts, in particular, greatly increased the use of difference-based heuristics and the extremeness of the data-based allocations. Fifth, the effects of graphical format were much greater than the effects of prior beliefs and temporal frame, as identified in prior research. Sixth and finally, neither real-world experience nor explicit training reduced these biases. We next address in greater detail three issues that we raised pre-

[11]The effects of semantic frame, graphic format, and prior beliefs were generally in the same directions as in the prior experiments, though only allocations to the local contrast variable were directly comparable because of the changed stimulus data. Only the effect of prior beliefs was statistically significant. Given the purpose of Experiment 3, we focused on the analyses that pooled data across experiments.

viously: extremeness in heuristic-based allocations, the causes of trend neglect, and what can be done to avoid the biases we have identified.

Extremeness in Heuristic-Based Allocations

To make predictions about the effects of the experimental manipulations, we relied on several parsimonious operationalizations of three focal types of heuristics. We summarize these predictions in Table 1, along with the mean allocations observed in our experiments. These means are mixtures of the distinct patterns of allocation associated with each type of heuristic. Thus, they suppress the extremeness in allocations that exists at the individual level. To provide a better summary measure of the patterns of allocation, Table 4 reports the mean allocations, conditioned on the variables chosen as most effective by each participant (i.e., the variable receiving the highest allocation). The choice of the most effective variable is useful because we designed the stimuli such that different heuristics predicted different choices. When the local contrast variable was chosen, the participant likely used a difference-based heuristic. When the global contrast variable was chosen in Experiments 1 and 2, our cumulative results suggest that most people used an exemplar-based heuristic, though a minority used a trend-based heuristic. In Experiment 3, choice of the control variable was a good indicator of the use of an exemplar-based heuristic, and choice of the global trend variable was a good indicator of the use of a trend-based heuristic. These mappings between allocations and heuristics were supported by the open-ended responses analyzed in Experiment 1.

Table 4 shows the conditional allocations for each graphic format and reveals several important insights. First, the allocation to the chosen variable is quite large (usually greater than 50 and at least 25 points greater than the next highest allocation) and is always considerably greater than the heuristic-based predictions in Table 1, which suggests that our operationalizations of these heuristics underestimated the amount of bias. Second, though seldom the most frequent heuristic, difference-based heuristics (i.e., choosing the local contrast variable) were usually the most extreme. This extremeness is especially evident for line charts, for which mean allocations to the local contrast variable were about 65, compared with 53 for the other formats. Third, for Experiment 3, in which trend-based heuristics uniquely favored the global trend variable, the proportion of participants choosing the global trend variable was small (i.e., trend neglect); however, allocations to that variable among those participants were large (M = 63). This pattern is consistent with our conclusion that line charts encourage both difference-based and trend-based heuristics. Fourth, we observe that the variable consistent with exemplar-based heuristics was generally chosen most frequently; however, the mean allocations to that variable were not as extreme as those for the other heuristics (ranging from 49 to 55). Overall, these results suggest the presence of two qualitatively different types of effects: the choice of a heuristic and the strength of the bias, given the use of a particular heuristic. Understanding the determinants of these effects is an important topic for further research.

The Causes of Trend Neglect

In our experiments, the participants used trend-based heuristics much less frequently than the difference-based and exemplar-based heuristics—as evidenced by the open-ended responses in Experiment 1 and the unpopularity of the global trend variable in Experiment 3.[12] Such trend neg-

[12]The Web Appendix (http://www.marketingpower.com/jmraug10) provides an analysis of allocations to the control variable in Experiment 2 that support trend neglect.

Table 4

SUMMARY ACROSS EXPERIMENTS OF MEAN ALLOCATIONS, CONDITIONED ON THE VARIABLE CHOSEN AS MOST EFFECTIVE

	Experiments 1 and 2 (N = 459)				*Experiment 3 (N = 160)*			
	Proportion	*Mean Allocations*			*Proportion*	*Mean Allocations*		
Variable Chosen as Most Effective	*Choosing*	*C*	*G*	*L*	*Choosing*	*C*	*G*	*L*
Bar Charts								
Control (C)	.02	**43**	32	25	.54[e]	**49**	23	29
Global trend (G)	.69[et]	21	**52**	27	.08[t]	17	**57**	27
Local contrast (L)	.22[d]	17	27	**55**	.28[d]	26	22	**53**
No unique maximum	.06	15	44	41	.10	40	20	40
Line Charts								
C	.05	**56**	21	24	.46[e]	**49**	25	26
G	.31[et]	21	**54**	25	.15[t]	13	**63**	24
L	.55[d]	14	19	**66**	.33[d]	21	17	**62**
No unique maximum	.09	30	33	37	.05	20	35	45
Tables								
C	.05	**52**	23	25	.37[e]	**55**	20	24
G	.57[et]	22	**51**	26	.17[t]	26	**50**	23
L	.32[d]	21	30	**50**	.30[d]	24	22	**54**
No unique maximum	.07	28	34	38	.16	34	31	35

Notes: For each participant, a variable was selected as most effective if it received the highest allocation (if two variables received the same highest value, neither was counted as having been selected, and it was assigned to the "No unique maximum" category). Bold values indicate the largest allocation in each condition.

[e]Allocations most consistent with exemplar-based heuristics.

[t]Allocations most consistent with trend-based heuristics.

[d]Allocations most consistent with difference-based heuristics.

[et]Allocations consistent with exemplar-based and trend-based heuristics; however, other results suggest that the majority used exemplar-based heuristics.

lect is consistent with Hutchinson and Alba's (1997) finding for tables, though it is nonetheless surprising that graphical formats did so little to overcome this bias, inasmuch as charts are often touted for their ability to make trends easier to apprehend. In Experiment 1, graphical formats increased the use of trend-based heuristics (e.g., self-reported use was 13% for line charts compared with 1% for tables). However, graphical formats had an even greater influence on the use of difference-based heuristics (e.g., 43% for line charts compared with 16% for tables), and difference-based heuristics were associated with stronger biases. Thus, the net effect was that trends still exerted relatively little influence on mean allocations. This result for simultaneously presented, multivariate information is even more surprising considering the trend effects often found for sequentially presented, univariate information (e.g., Zauberman, Diehl, and Ariely 2006) and the accurate perception of bivariate linear and nonlinear trends in various charts and scatter plots (e.g., Best, Smith, and Stubbs 2007; Doherty et al. 2007).

Why might overall trends be neglected? A well-known visual principle works against trend-based and for difference-based heuristics. The eye analyzes small regions in detail, jumping between regions; the brain constructs the whole picture on the basis of locally determined parts, which are assembled into objects and then assembled into whole scenes (e.g., Biederman 1987; Goertz and Goertz 2004). This principle is nicely illustrated with impossible figures such as the devil's trident in Figure 5 (Panel A). The left side is locally perceived as three pipes, and the right side is locally perceived as a block solid. However, the two regions cannot be reconciled as a single physical object, because the middle pipe is disconnected from the surfaces of the block. The paradox is initially difficult to detect, because all local cues are consistent with some physically possible object, and only a global mapping of the local surfaces reveals the impossibility of the whole object.

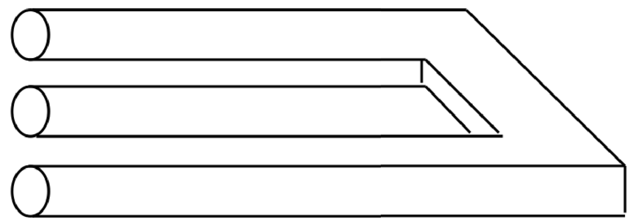
There is a similar local dominance in our global trend and local contrast variables, as Panel B of Figure 5 illustrates. Visually, X_{Local1} better matches Y_1 than does $X_{Global1}$. Similar to Y_1, X_{Local1} is concave and increasing on the left side, where $X_{Global1}$ is linearly increasing. In the middle, Y_1 and X_{Local1} exhibit matching peaks and valleys, and $X_{Global1}$ is still linearly increasing. On the right side, Y_1 and X_{Local1} are both decreasing, where $X_{Global1}$ is still linearly increasing. In actuality, Y_1 is the sum of X_{Local1} and $X_{Global1}$, and $X_{Global1}$ has a higher simple and partial correlation with Y_1 than does X_{Local1}. Moreover, Y_1 is less curved and more strongly increasing than it should be were it linearly related only to X_{Local1}. These global properties are difficult to detect visually. To appreciate them, note that Y_1' is perfectly linearly related to only X_{Local1}, revealing that Y_1 is too flat, relative to its overall increasing slope, to be related only to X_{Local1}. Though supported by different cognitive processes, this general principle of local dominance also applies to tables. It is relatively easy to do the mental arithmetic for adjacent differences but mentally difficult to "smooth" the numbers for each variable into some general trend.

For other types of visual stimuli, global dominance occurs instead of local dominance. For example, when a large letter "R" is formed out of many, physically smaller letters "H," the large letter is easier to detect than the small

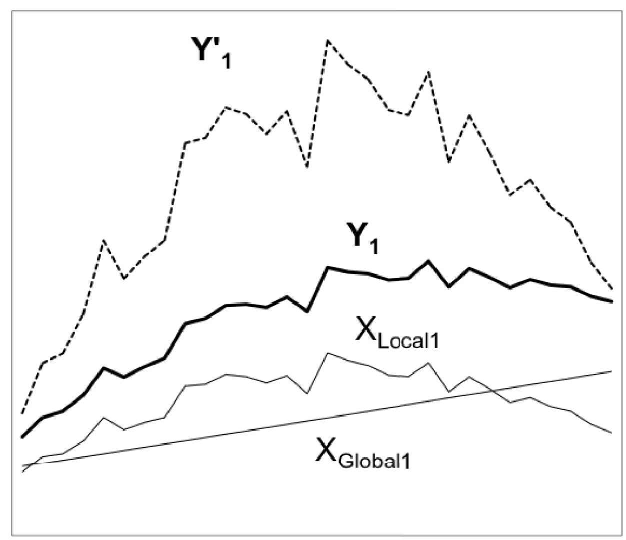
Figure 5

A COMPARISON OF IMPOSSIBLE FIGURES AND GLOBAL TREND AND LOCAL CONTRAST CURVES THAT ILLUSTRATES THE PERCEPTUAL NATURE OF TREND NEGLECT FOR LINE CHARTS

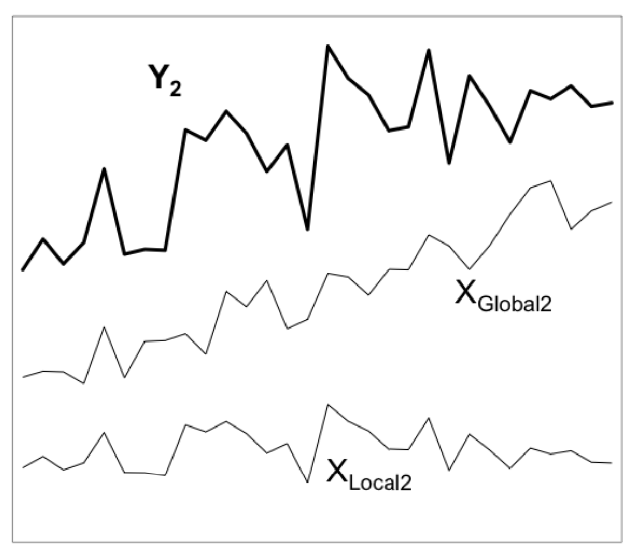
A: Impossible Figure Designed by Penrose



B: An Example of Local Dominance



C: An Example of Global Dominance



letters (Navon 2003). This type of hierarchical nesting of patterns inspired the curves in Panel C of Figure 5, where the match between the small deviations in Y_2 and X_{Local2} are more difficult to detect than those in Panel B, but the match in the general trend of Y_2 and $X_{Global2}$ is easy to detect (though both the simple and partial correlations are larger for Y_2 and X_{Local2} than for Y_2 and $X_{Global2}$). The connection between visual perceptions of objects and the visual analysis of statistical charts appears to be a promising direction for further research.

What to Do?

The results reported herein demonstrate that data-based inferences are subject to strong, heuristic-based biases that are not reduced by graphical presentations of the data, real-world experience, or explicit training. What recommendations can we offer to avoid such biases? There seems to be only one obvious solution: Always estimate some statistical model when displaying multivariate data (e.g., linear regression), especially when the intent is to show causal or correlational relationships. Even when the sample size is very small, and the confidence intervals of the estimated coefficients are very large, the coefficients are nonetheless unbiased in their description of the data. The interpretation of the coefficients is still a matter of human reasoning, but at least their magnitudes will be accurate. Consistent with this recommendation, Wainer and Velleman (2001, p. 332) conclude their section on multivariate graphics by saying, "displays must be integrated with analyses so that the data analyst can move smoothly from looking at aspects of the data to quantitative descriptions and tests and then back again to examine residuals or look for additional patterns." It is also of interest that managers in Experiment 2 complained much more than students about their inability to conduct a statistical analysis. Thus, the "graph-plus-model" guideline may be well received for business data. We also note that Tufte (2001, 2006), arguably the most influential commentator on the graphic display of data, has frequently maintained that data analysis must precede data display because display is a mainly communication tool.

In conclusion, although charts and tables have always been a part of business analyses, the increased penetration of personal computers, spreadsheets, and user-friendly statistical software creates the risk of making these tools, once reserved for sophisticated analysts, into "decision traps" for the unwary manager. This research identifies some of those traps and the types of decision-making heuristics that create them. We hope research will extend this approach to a broader array of data-based decisions and data display formats.

REFERENCES

Alba, Joseph W. and J. Wesley Hutchinson (2000), "Knowledge Calibration: What Consumers Know and What They Think They Know," *Journal of Consumer Research*, 27 (September), 123–56.

Arkes, Hal R. (1991), "Cost and Benefits of Judgment Errors: Implications for Debiasing," *Psychological Bulletin*, 110 (November), 486–98.

Benbasat, Izak and A.S. Dexter (1985), "An Experimental Evaluation of Graphical and Color-Enhanced Information Presentation," *Management Science*, 31 (November), 1348–64.

Best, Lisa A., Laurence D. Smith, and D. Alan Stubbs (2007), "Perception of Linear and Nonlinear Trends: Using Slope and Curvature Information to Make Trend Discriminations," *Perceptual and Motor Skills*, 104 (June), 707–721.

Biederman, I. (1987), "Recognition-by-Components," *Psychological Review*, 94 (2), 115–47.

Biyalogorsky, Eyal, W. Boulding, and Richard Staelin (2006), "Stuck in the Past: Why Managers Persist with New Product Failures," *Journal of Marketing*, 70 (April), 108–121.

Bolton, Lisa E. (2003), "Stickier Priors: The Effects of Nonanalytic Versus Analytic Thinking in New Product Forecasting," *Journal of Marketing Research*, 40 (February), 65–79.

Boulding, William, Ruskin Morgan, and Richard Staelin (1997), "Pulling the Plug to Stop the New Product Drain," *Journal of Marketing Research*, 34 (February), 164–76.

Broniarczyk, S.M. and J.W. Alba (1994), "Theory Versus Data in Prediction and Correlation Tasks," *Organizational Behavior and Human Decision Processes*, 57 (May), 117–39.

Carpenter, P.A. and P. Shah (1998), "A Model of the Perceptual and Conceptual Processes in Graph Comprehension," *Journal of Experimental Psychology: Applied*, 4 (June), 75–100.

Chapman, Loren J. and J.P. Chapman (1969), "Illusory Correlation as an Obstacle to the Use of Valid Psychodiagnostic Signs," *Journal of Abnormal Psychology*, 74 (June), 271–80.

Chua, Hannah Faye, J. Frank Yates, and Priti Shah (2006), "Risk Avoidance: Graphs Versus Numbers," *Memory & Cognition*, 34 (March), 399–410.

Diehl, Kristin, Laura J. Kornish, and John G. Lynch Jr. (2003), "Smart Agents: When Lower Search Costs for Quality Information Increase Price Sensitivity," *Journal of Consumer Research*, 30 (June), 56–71.

Doherty, Michael E., Richard B. Anderson, Andrea M. Angott, and Dale S. Klopfer (2007), "The Perception of Scatterplots," *Perception & Psychophysics*, 69 (October), 1261–72.

Fong, G.T., David H. Krantz, and Richard E. Nisbett (1986), "The Effects of Statistical Training on Thinking About Everyday Problems," *Cognitive Psychology*, 18 (July), 253–92.

Goertz, Ralf and Claudia Goertz (2004), "Hemispheric Differences in the Recognition of Possible and Impossible Objects," *Laterality*, 9 (2), 209–24.

Harvey, N. and Fergus Bolger (1996), "Graphs Versus Tables: Effects of Data Presentation Format on Judgmental Forecasting," *International Journal of Forecasting*, 12 (March), 119–37.

Henry, Gary T. (1995), *Graphing Data: Techniques for Display and Analysis*. Thousand Oaks CA: Sage Publications.

Hutchinson, J. Wesley and Joseph W. Alba (1997), "Heuristics and Biases in the 'Eye-Balling' of Data: The Effects of Context on Intuitive Correlation Assessment," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23 (May), 591–621.

———, Wagner A. Kamakura, and John G. Lynch Jr. (2000), "Unobserved Heterogeneity as an Explanation for 'Reversal' Effects in Behavioral Research," *Journal of Consumer Research*, 27 (December), 324–44.

Jarvenpaa, Sirkka L. (1990), "Graphic Displays in Decision Making: The Visual Salience Effect," *Journal of Behavioral Decision Making*, 3 (October–December), 247–62.

Juslin, Peter, Henrik Olsson, and A.C. Olsson (2003), "Exemplar Effects in Categorization and Multiple-Cue Judgment," *Journal of Experimental Psychology: General*, 132 (1), 133–56.

Kahn, Barbara E., Mary Frances Luce, and Stephen M. Nowlis (2006), "Debiasing Insights from Process Tests," *Journal of Consumer Research*, 33 (June), 131–38.

Lane, David M., Craig A. Anderson, and Kathryn L. Kellam (1985), "Judging the Relatedness of Variables: The Psychophysics of Covariation Detection," *Journal of Experimental Psychology: Human Perception and Performance*, 11 (October), 640–49.

Larrick, R.P. (2004), "Debiasing," in *Blackwell Handbook of Judgment and Decision Making*, D.J. Koehler and N. Harvey, eds. Malden, MA: Blackwell Publishing, 316–37.

Lurie, Nicholas H. and Charlotte H. Mason (2007), "Visual Representation: Implications for Decision Making," *Journal of Marketing*, 71 (January), 160–77.

Mantrala, Murali K., Prabhakant Sinha, and Andris A. Zoltners (1992), "Impact of Resource Allocation Rules on Marketing Investment-Level Decisions and Profitability," *Journal of Marketing Research*, 23 (May), 162–75.

Meyer, Joachim, Marcia Kuskin Shamo, and Daniel Gopher (1999), "Information Structure and the Relative Efficacy of Tables and Graphs," *Human Factors*, 41 (December), 570–87.

Meyer, Robert (1989), "A Theory of the Inductive Learning of Multiattribute Preferences," in *Methodological Advances in Consumer Research*, David Brinberg and Richard Lutz, eds. Berlin: Springer-Verlag, 253–85.

Navon, David (2003), "What Does a Compound Letter Tell the Psychologist's Mind?" *Acta Psychologica*, 114 (3), 273–309.

Nisbett, R.E., D.H. Krantz, C. Jepson, and Z. Kunda (1983), "The Use of Statistical Heuristics in Everyday Inductive Reasoning," *Psychological Review*, 90 (October), 339–63.

Simkin, David and Reid Hastie (1987), "An Information-Processing Analysis of Graph Perception," *Journal of the American Statistical Association*, 82 (June), 454–65.

Smith, Laurence D., Lisa A. Best, D. Alan Stubbs, Andrea Bastiani Archibald, and Roxann Roberson-Nay (2002), "Constructing Knowledge: The Role of Graphs and Tables in Hard and Soft Psychology," *American Psychologist*, 57 (October), 749–61.

Tufte, Edward R. (2001), *Visual Display of Quantitative Information*, 2d ed. Cheshire, CT: Graphics Press.

——— (2006), *Beautiful Evidence*. Cheshire, CT: Graphics Press.

Van Osselaer, Stijn M.J., Chris Janiszewski, and Marcus Cunha Jr. (2004), "Stimulus Generalization in Two Associative Learning Processes," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30 (3), 626–38.

Wainer, Howard and Paul F. Velleman (2001), "Statistical Graphics: Mapping the Pathways of Science," in *Annual Review of Psychology*, Vol. 52, S.T. Fiske, D.L. Schacter, and C. Zahn-Waxler, eds. Palo Alto, CA: Annual Reviews, 305–335.

Wallgren, Anders, Britt Wallgren, Rolf Persson, Ulf Jorner, and Jan-Aage Haaland (1996), *Graphing Statistics and Data: Creating Better Charts*. Newbury Park, CA: Sage Publications.

Zacks, Jeff and Barbara Tversky (1999), "Bars and Lines: A Study of Graphic Communication," *Memory & Cognition*, 27 (November), 1073–1079.

Zauberman, Gal, Kristin Diehl, and Dan Ariely (2006), "Hedonic Versus Informational Evaluations: Task Dependent Preferences for Sequences of Outcomes," *Journal of Behavioral Decision Making*, 19 (July), 191–211.